

Ella Mustanoja

# Ribosomaalisen RNA:n poistomenetelmän kehittäminen suolistomikrobinäytteistä tehtävän cDNA-sekvensointikirjaston valmistusta varten

Metropolia Ammattikorkeakoulu

Laboratorioanalyttikko (AMK)

Laboratorioanalytiikka

Opinnäytetyö

7.5.2018

Tekijä Otsikko  Sivumäärä Aika	Ella Mustanoja Ribosomaalisen RNA:n poistomenetelmän kehittäminen suolistomikrobinäytteistä tehtävän cDNA-sekvensointikirjaston valmistusta varten 24 sivua + 1 liite 7.5.2018
Tutkinto	Laboratorioanalyttikko (AMK)
Tutkinto-ohjelma	Laboratorioanalytiikka
Ammatillinen pääaine	
Ohjaajat	Tutkimusjohtaja Petri Auvinen Laboratorioinsinööri Lars Paulin Lehtori Tiina Soininen
<p>Tämä opinnäytetyö tehtiin Biotekniikan instituutille. Opinnäytetyön tavoitteena oli luoda uusi menetelmä ribosomaalisen RNA:n poistamiseksi RNA-näytteestä suolistomikrobiston aktiivisten geenien tutkimusta varten. Geenien aktiivisuutta voidaan tutkia näytteen lähetti-RNA:ta sekvensoimalla, ja tätä tutkimuksen alaa kutsutaan metatranskriptomiikaksi. Tätä varten tulee näytteistä poistaa tai vähentää ribosomaalista RNA:ta, koska lähetti-RNA:n osuus RNA:sta on vain 1–5 %. Ilman rRNA:n poistoa tehdystä sekvensoinnista saatava data sisältäisi siis yli 90 % tutkimukselle turhaa tietoa.</p> <p>Eukaryootti mRNA:ssa olevan polyadenosiinihännän vuoksi eukaryoottinen mRNA on mahdollista eristää muusta RNA:sta oligo(dT)-selektiolla. Prokaryoottisolujen kohdalla täytyy turvautua rRNA:n poistamiseen tai vähentämiseen näytteestä, sillä mRNA:ta ei pystytä eristämään muusta RNA:sta.</p> <p>Menetelmän kehitykseen käytettiin kolmea maitohappobakteerikantaa. Kullekin bakteerikannalle amplifioitiin rRNA:lle komplementaariset biotinyloidut yksijuosteiset PCR-fragmentit, jotka hybridisoitiin RNA-näytteeseen, ja PCR-fragmenteihin hybridisoitu rRNA digestoitui entsyymaattisesti. Tämän jälkeen näytteistä valmistettiin cDNA-kirjastot kaupallisella kitillä ja kirjastot sekvensoitiin.</p> <p>Näytteiden sekvensseistä 55–76 % mappautui rRNA-operoniin. Sekvenssit, jotka mappautuivat rRNA:han, olivat peräisin PCR-fragmenttien päiden alueilta. Menetelmällä saatiin siis rRNA:ta osin poistettua, mutta sinällään se ei kuitenkaan vielä ole käyttökelpoinen, joten menetelmän kehitystä tulee jatkaa. Kun menetelmä saadaan toimimaan testissä käytetyille testikannoille, menetelmää kokeillaan suolistomikrobinäytteille. Yksittäisestä suolistomikrobinäytteestä valmistetaan biotinyloidut PCR-fragmentit universaaleja alukkeita käyttäen, jolloin saadaan kaikkia näytteessä olevia bakteereja vastaavia PCR-fragmenteja, ja näitä fragmentteja avuksi käyttäen rRNA:n osuutta RNA-näytteessä voidaan vähentää samaa menetelmää hyväksikäyttäen.</p>	
Avainsanat	rRNA, cDNA-sekvensointi, sekvensointikirjasto

Author Title Number of Pages Date	Ella Mustanoja Developing a Method to Remove Ribosomal RNA for Preparation of a cDNA seq Library from Gut Microbe Samples 24 pages + 1 appendix 7 May 2018
Degree	Bachelor of Laboratory Services
Degree Programme	Laboratory sciences
Professional Major	
Instructors	Petri Auvinen, Research Director Lars Paulin, Laboratory engineer Tiina Soininen, Senior Lecturer
<p>This study was done at the Institute of Biotechnology. The aim of the project was to develop a method for decreasing the amount of ribosomal RNA in an RNA sample, which is then used in a cDNA library preparation process. The method will be used for gut microbiota research. Active genes in gut microbiota can be studied using metatranscriptomics where messenger RNA is sequenced. Only 1-5 % of the total RNA is messenger RNA and most of the RNA is ribosomal RNA. This is why sequencing without rRNA depletion generates over 90 % of useless information. This is why rRNA depletion is crucial.</p> <p>Because of a polyA-tail in eukaryotic mRNA, eukaryotic mRNA can be separated from the rest of the RNA using oligo(dT) selection. Prokaryotic mRNA doesn't include any specific feature to separate it from the rest of the RNAs, so methods of decreasing the amount of rRNA have been developed.</p> <p>In this method development project three lactic acid bacteria strains were used as test samples. In this method, biotinylated single-stranded PCR fragments complementary to rRNA were amplified for each bacterium. These fragments were then hybridized to RNA-samples and the RNA, which was hybridized to the fragments, was then enzymatically digested. cDNA libraries were prepared from these samples with a commercial kit and the libraries were sequenced.</p> <p>55 to 76 % of the sequences were mapped to rRNA operon. Reads that were mapped to the operon were from certain areas from the ends of the used PCR fragments. So, some of the rRNA was successfully depleted by using the method. The method, as it is now, isn't quite usable but needs further development. When the method works with the test samples, it can be tested with the gut microbe samples. PCR fragments are then amplified using universal primers and one gut microbiota DNA sample. This way it is possible to amplify PCR-fragments for all the bacterial species in the gut sample and the amount of rRNA can be decreased, using the same method as developed for the lactic acid bacteria.</p>	
Keywords	rRNA, cDNA-sequencing, sequencing library

# Sisällys

## Lyhenteet

1	Johdanto	1
2	Metatranskriptomiikka	2
2.1	Genomiikka	2
2.2	Sekvensointi	2
2.3	RNA:n muokkaus sekvensointitarkoituksiin	5
2.3.1	RNA:n eristys	5
2.3.2	Ribosomaalisen RNA:n poistaminen RNA-näytteestä	6
2.3.3	cDNA-sekvensointikirjaston valmistus	7
2.3.4	Sekvensointikirjaston laadun varmistaminen	8
3	Materiaalit ja menetelmät	8
3.1	Käytetyt materiaalit	8
3.2	Menetelmät	9
3.2.1	PCR-fragmenttien amplifiointi	9
3.2.2	PCR-fragmenttien puhdistus	13
3.2.3	PCR-fragmenttien sitominen streptavidiini-magneettihelmille	14
3.2.4	dsDNA:n muokkaus ssDNA:ksi	14
3.2.5	rRNA:n poistaminen RNA-näytteestä	14
3.2.6	RNA-sekvensointikirjaston valmistus	15
4	Tulokset	17
4.1	PCR-fragmenttien valmistus	17
4.2	Sekvensointikirjaston valmistus	17
4.3	Sekvensoinnin tulokset	19
	Yhteenveto	23
	Lähteet	24
	Liitteet	
	Liite 1. PCR-fragmenttien agaroosigeelielektroforeesikuvat	

## 1 Johdanto

Opinnäytetyö tehtiin Helsingin yliopistossa Biotekniikan instituutin DNA sekvensointi- ja genomiikkalaboratoriossa. Laboratoriossa tehdään muun muassa Parkinsonin tautiin liittyvää tutkimusta. Ryhmässä on tutkittu Parkinson potilaiden suoliston, nenän ja suun mikrobistoa 16S rRNA -geeniä sekvensoimalla. Tutkimuksissa on löytynyt eroa suoliston ja suun mikrobistossa Parkinson potilaiden ja kontrolliryhmän välillä. *Prevotellaceae* -heimon mikrobien määrä Parkinson potilaiden ulosteessa oli vähentynyt yli 70 % verrattuna kontrolliryhmään (Scheperjans ym. 2015). Suun mikrobistossa eroa oli joidenkin yksittäisten bakteeriheimojen ja -sukujen kohdalla, mikä oli nähtävissä sekä vähentyneinä että lisääntyneinä bakteerimäärinä (Pereira ym. 2017). Ryhmän tekemän tutkimuksen tavoitteena on selvittää, voidaanko mikrobistosta löytää biomarkkereita, joiden avulla Parkinsonin tauti voitaisiin todeta varhaisessa vaiheessa ennen motoristen oireiden alkua. Tähän mennessä tehty tutkimus ei todista mikrobiston ja Parkinsonin taudin välisiä syyseuraussuhteita, vaan sen, että mikrobistoissa on eroa. Seuraavassa tutkimuksen vaiheessa on tarkoitus kartoittaa metatranskriptomiikan avulla mikrobiston geenien aktiivisuutta.

Aiemmin kerätyistä ja eristetyistä näytteistä on tehty shotgun-metagenomikirjasto, josta mikrobiston eroja on voitu tutkia. Geenien aktiivisuutta voidaan tutkia cDNA-sekvensoinnilla, jota varten tulee näytteistä valmistaa cDNA-sekvensointikirjastoja. cDNA-kirjasto valmistetaan RNA-näytteestä, joka muutetaan käänteistranskriptiolla komplementaariseksi DNA:ksi (cDNA). cDNA-sekvensointikirjaston valmistus vaatii ribosomaalisen RNA:n (rRNA) poistamista RNA-näytteistä jossakin sekvensointikirjaston valmistuksen vaiheessa, sillä rRNA:n osuus on yli 85 % koko solun RNA:sta, ja näin ollen sekvensoinnista saatavasta datasta valtaosa olisi tutkimuksen kannalta tarpeetonta tietoa (Petrova ym. 2017). Kaupalliset kitit prokaryootti rRNA:n poistamiseksi ovat huomattavan kalliita, joten laboratoriolla oli suuri tarve oman menetelmän kehittämiseksi. Tämän opinnäytetyön tavoitteena oli kehittää menetelmä rRNA:n poistamiseen RNA-näytteestä ennen cDNA-sekvensointikirjaston valmistusta. Menetelmän kehitysvaiheessa käytettiin kolmea maitohappobakteeria testinäytteinä.

## 2 Metatranskriptomiikka

### 2.1 Genomiikka

Genomiikka tarkoittaa eliön perintöaineksen eli genomin kartoitusta, sekvensointia ja analysointia. Metagenomiikka on tietystä ympäristöstä saatavan geneettisen materiaalin analysointia. Transkriptomiikka on transkriptomin, eli transkriptoituneen lähetti-RNA:n eli mRNA:n analysointia ja vastaavasti metatranskriptomiikka tietyssä ympäristössä olevan mRNA:n analysointia. (Suominen ym. 2010.)

Sekvensoinnista saatavan tiedon analysointia kutsutaan bioinformatiikaksi. Bioinformatiikka käsitteenä voidaan selittää usealla tavalla. Alan nopea kehitys on vaikeuttanut yksiselitteisen määritelmän käyttämistä. Bioinformatiikka on biotieteiden ja informaatiotieteiden yhdistelmä, ja sitä käytetään biologisten makromolekyylien tietokoneavusteiseen analysointiin. Bioinformatiikka voidaan nähdä alana, joka kehittää informaatio- ja tietoteknisiä välineitä biologisten kysymysten tutkimiseen ja toisaalta se voidaan nähdä alana, jota käytetään biologisen informaation tallentamiseen ja analysoimiseen.

Bioinformatiikkaa on kehitetty sekvensointimenetelmien kehittymisen seurauksena räjähdysmäisesti kasvaneen biologisen informaation käsittelyn tarpeeseen. Biotietokannat sisältävät tietoa makromolekyylien (DNA, RNA, proteiinit) sekvensseistä sekä niihin liittyvää muuta informaatiota, kuten tunnistettuja geenejä.

### 2.2 Sekvensointi

Sekvensointia tehdään nukleiinihappojen emäsjärjestyksen määrittämiseksi. Ensimmäiset sekvensointimenetelmät on kehitetty 1970-luvun puolivälin jälkeen. Ensimmäinen ja aina viime vuosiin käytetyin sekvensointimenetelmä on Sangerin menetelmä. Sangerin menetelmä perustuu DNA:n synteisiin DNA-juostetta templaattina käyttäen ja dideoksinukleotidien käyttöön. Kun dideoksinukleotideja on vähän suhteessa deoksinukleotideihin, DNA:n syntetisointi loppuu sattumanvaraisesti johonkin dideoksinukleotidiin, jolloin synteesituotteiksi saadaan useita erimittaisia DNA-nauhoja. Nämä syntyneet DNA-nauhat voidaan erotella toisistaan koon perusteella kapillaarielektroforeesilla. Kun dideoksinukleotidit on leimattu, kukin nukleotidi omalla leimallaan, voidaan nukleotidijärjestys määrittää kapillaarielektroforeesilla. (Suominen ym. 2010.)

Sekvensointimenetelmät ovat kehittyneet paljon, minkä seurauksena sekvensoinnista on tullut huomattavasti tehokkaampaa: saatavan datan määrä on kasvanut räjähdysmäisesti ja kustannukset ovat laskeneet. Näitä sekvensointimenetelmiä kutsutaan uuden sukupolven sekvensointimenetelmiksi (NGS, engl. next-generation sequencing). NGS-menetelmien myötä biologisen informaation määrä on kasvanut valtavasti viimeisen vuosikymmenen aikana.

NGS-menetelmät voidaan karkeasti jakaa synteesisekvensointitekniikoihin (engl. sequencing by synthesis) ja ligaatiosekvensointitekniikoihin (engl. sequencing by ligation). Sangerin menetelmä perustuu synteesisekvensointiin. Sangerin menetelmässä ja muissa synteesisekvensointimenetelmissä sekvensoitavalle templaattijuosteelle sekvensoidaan vastinjuostetta DNA-polymeraasin avulla. NGS-menetelmissä kiinteällä kantajapinnalla on immobilisoituneena DNA-adapteri, johon DNA-templaatti kiinnitetään. Tämän vuoksi sekvensointikirjaston valmistuksen yhteydessä DNA-templaattiin kiinnitetään adapteri, joka on sekvensointialustan DNA-adapterille komplementaarinen (Mardis, 2017). NGS-menetelmistä pyrosekvensointi eli 454-sekvensointimenetelmä perustuu synteesisekvensointiin. Pyrosekvensoinnissa DNA-templaatin kiinnittämisen jälkeen lisätään yksitellen kutakin nukleotiditrifosfaattia. Uusi fosfodiesterisidos syntyy, kun reaktioon lisätään templaatti-DNA:lle komplementaarista nukleotidia. Kun DNA-polymeraasi liittää syntetisoitavaan juosteeseen uuden nukleotidin, dNTP:stä irtaantuu yksi pyrofosfaattimolekyyli, josta ATP-sulfurylaasientsyymi muodostaa ATP:tä. Lusiferiasientsyymi katalysoi lusiferiinin hapettumista oksilusiferiiniksi (Suominen ym. 2010). Tässä reaktiossa syntyy valoa, joka voidaan detektoida. Syntyvän valon intensiteetti on verrannollinen sitoutuneiden nukleotidien määrään. (Stranneheim & Lundeberg, 2012.)

Toinen synteesisekvensointimenetelmä perustuu palautuvien värillisten terminaattorinukleotidien käyttöön. Tähän menetelmään perustuu Illumina-sekvensointi. Tekniikassa DNA-templaatti hybridisoidaan tasomaiselle pinnalle, jolla jokainen templaatti amplifoidaan käyttäen silta-amplifiointia. Tällöin syntyy tiheitä klustereita, joissa jokaisessa on yksilöllinen DNA-templaatti. Amplifioinnin jälkeen klusterit sekvensoidaan käyttäen palautuvia terminaattorinukleotideja. Yhdessä reaktiossa tapahtuu vain yhden emäksen ekstensio, joka voidaan detektoida fluoresenssikuvantamisella. Tämän jälkeen väri ja terminaattori poistetaan kemiallisesti, minkä jälkeen uusi sekvensointireaktio voidaan aloittaa. (Stranneheim & Lundeberg, 2012.)

Ligaatiosekvensointi perustuu DNA-ligaasientsyymin herkkyyteen väärälle emäspariutumiselle. Solid-sekvensointi perustuu ligaatiosekvensointiin. Sekvensoitavan tuntemattoman DNA-alueen vieressä tulee olla tunnettu DNA-alue. Lyhyt komplementaarinen ankkuri-DNA-jakso kiinnitetään tunnetulle alueelle. Seuraavaksi lisätään koettimia, jotka tyypillisesti 8 nukleotidin mittaisia oligonukleotideja, jotka on leimattu erilaisilla fluoresoivilla merkkiaineilla sen mukaan, mitkä kaksi nukleotidia on oligonukleotidin 3'-päässä. Koetinseoksen oligonukleotidit koostuvat seuraavasti: Kaksi ensimmäistä 3'-pään nukleotidia on tunnettuja. Kolme seuraavaa nukleotidia ovat kaikkina mahdollisina yhdistelminä. Näiden nukleotidien jälkeen on katkaisukohta, jonka jälkeen on kolme neutraalia emästä, jotka voivat hybridisoitua mihin tahansa nukleotidiin. Merkkiaine on kiinni oligonukleotidin 5'-päässä. Näistä koettimista sellainen, joka on komplementaarinen sekvensoitavalle DNA:lle ankkuri-DNA:sta eteenpäin, hybridisoituu templaattiin ja ligoituu ankkuri-DNA:han. Ylimääräiset koettimet huuhdotaan pois ja sitoutuneen koettimen fluoresenssivalo voidaan detektoida ja päätellä, mitkä nukleotidit kyseisessä kohdassa on. Fluoresoiva merkkiaine poistetaan katkaisemalla oligonukleotidi, minkä jälkeen ligaatio ja detektointi toistetaan useita kertoja. Näin saadaan sekvenssi, jossa toistuu kahden tunnetun ja kolmen tuntemattoman nukleotidin jaksot. Tämän jälkeen menettely toistetaan lyhentämällä ankkuri-DNA:ta aina yhdellä nukleotidilla, ja lopulta kaikki sekvenssin nukleotidit on saatu luettua kahdesti. (Stranneheim & Lundeberg, 2012.)

Taulukkoon 1 on koostettu sekvensointialustojen sekvensointiperiaatteet, saatavien sekvenssien pituudet ja sekvensointialustojen käyttökohteet.



Taulukko 1. Sekvensointialustojen sekvensointiperiaatteet ja käyttökohteet sekä saatavien readien pituudet. (Muokattu Mardis, 2017 ja Stranneheim & Lundeborg, 2012 pohjalta)

Sekvensointialusta	Sekvensointiperiaate	Sekvenssin max pituus (bp)	Käyttökohteet
454/Roche	Pyrosekvensointi	400-1000	Bakteeri- ja virusgenomit, multiplex-PCR tuotteet, pistemutaatioiden validointi ja targetoitujen somaattisten mutaatioiden detektointi
Illumina	Synteesisekvensointi	300+300	Monimutkaiset genomit, RNA-sekvensointi
ABI SOLiD	Ligaatiosekvensointi	75+35	Monimutkaiset genomit, RNA-sekvensointi
Pacific Biosciences	Single molecule sekvensointi	< 100 000	Monimutkaiset genomit, mikrobiologiset ja infektio-tauti genomit, metylaatioiden detektio
Ion Torrent	Ionisensiitiivinen synteesisekvensointi	200-400	Multiplex-PCR tuotteet,
Oxford Nanopore	Single molecule sekvensointi	Jopa satoja tuhansia	Patogeenivalvonta
Qiagen GeneReader	Synteesisekvensointi	107	Targetoidut mutaatio detektiot, syöpien nestebiopsiat

Tässä opinnäytetyössä käytettiin Sanger-sekvensointia PCR-fragmenttien sekvenssien varmistamiseen. Lisäksi varsinaiset cDNA-kirjastot sekvensoitiin käyttäen Illuminan Miseq -sekvensointialustaa. Illuminan Miseq -sekvensointi on synteesisekvensointitekniikka.

## 2.3 RNA:n muokkaus sekvensointitarkoituksiin

### 2.3.1 RNA:n eristys

RNA eli ribonukleiinihappo rakentuu toisiinsa liittyneistä nukleotideista, jotka koostuvat kolmesta osasta: fosfaattiryhmä, sokeriosana oleva riboosi ja orgaaninen emäs. Transkriptomiikkaa ajatellen oleellinen eristettävä RNA on mRNA, joka on hyvin nopeasti hajoavaa. mRNA:n elinikä on vain muutamia minutteja. Solujen omat ribonukleasientsyymit (RNAasit) sekä ympäristössä olevat RNAasit pilkkovat RNA:ta, minkä vuoksi ehjien ja täyspitkien RNA-molekyylien eristäminen on hankalaa. Kaikki eristyksessä käytettävät välineet tulee olla vapaita RNAasista ja RNAasi-inhibiittorien käyttö on tarpeen tullen suositeltavaa. (Suominen ym. 2010.)

Bakteerisoluista RNA:n eristuksen ensimmäinen vaihe, solujen hajottaminen voidaan tehdä useilla vaihtoehtoisilla menetelmillä lajista ja lähtötilanteesta riippuen. Hajotus voidaan tehdä mm. detergentin avulla tai ultraäänikäsittelyllä. Hajotuksen jälkeen näytteessä olevat proteiinit poistetaan. Aiemmin proteiinien poistoon käytettiin fenoli-kloroformiuuttoja, mutta nykyään käytetään silikakantajaperusteisia menetelmiä. Näytteessä oleva DNA voidaan poistaa esimerkiksi DNAasi I -käsittelyllä tai RNA:n selektiivisellä LiCl-saostuksella. (Suominen ym. 2010.)

### 2.3.2 Ribosomaalisen RNA:n poistaminen RNA-näytteestä

Bakteerin transkriptomi koostuu mm. proteiinia koodaavasta RNA:sta, tRNA:sta, tmRNA:sta (engl. transfer messenger RNA), sRNA:sta (engl. small regulatory RNA) ja rRNA:sta. Näistä rRNA:n osuus on yli 85 % koko solun RNA:sta. (Petrova ym. 2017.) Transkriptomia sekvensoitaessa suurimpia haasteita on mRNA:n kohtalaisen pieni osuus (1–5 %) kokonais-RNA:sta (He ym. 2010). Eukaryootti mRNA:ssa olevan polyadenosiinihännän vuoksi eukaryoottinen mRNA on mahdollista eristää muusta RNA:sta oligo(dT)-selektiolla. Prokaryoottisolujen kohdalla täytyy turvautua rRNA:n poistamiseen tai vähentämiseen näytteestä, sillä mRNA:ta ei pystytä eristämään muusta RNA:sta.

Ribosomaalisen RNA:n poistamiseksi on kehitetty useita menetelmiä ja markkinoilla on valmiita kaupallisia kittejä. Näissä kiteissä käytettäviä menetelmiä ovat mm. vähentävä hybridisaatio rRNA-spesifeillä koettimilla, rRNA:han erityisesti toimivalla eksonukleaa-silla digestointi ja poly(A)-hännän lisääminen rRNA:han, jolloin voidaan käyttää oligo(dT)-selektiota. Lisäksi rRNA:n poistaminen voidaan tehdä käänteistranskriptiolla rRNA-spesifeillä alukkeilla, jonka jälkeen RNAasi H-entsyymillä voidaan hajottaa rRNA-DNA-hybridit. Myös kokoerottelua geelielektroforeesilla voidaan käyttää rRNA:n poistossa niin, että valitaan koon mukaan vyöhykkeet, jotka eivät ole rRNA:ta. Kaupallisten kittien toimivuus on hyvin mikrobisidonnaista ja riippuu myös RNA:n laadusta. (He ym. 2010.)

Morlan ym. (2012) kehittivät menetelmän rRNA:n selektiiviseen poistoon. Menetelmässä 50–80 emäksen mittaiset antisense-DNA koettimet ovat komplemetaarisia poistettavan alueen kanssa. Nämä sekvenssit muodostavat RNA-DNA-hybridejä koko poistettavalle alueelle, myös fragmentoituneille alueille. RNAasi H -entsyymi pilkkoo RNA-DNA-hybridit ja DNAasi I poistaa ylimääräiset koettimet. Tässä menetelmässä on haas-

teena metatranskriptomiikan kannalta se, että mikrobien sekvenssit rRNA:ssa vaihtelevat lajikohtaisesti, joten tarvitaan valtava määrä lajispesifejä koettimia.

### 2.3.3 cDNA-sekvensointikirjaston valmistus

cDNA-sekvensointikirjasto on sekvensointia varten valmistettu näyte, joka sisältää yksitai kaksijuosteista cDNA:ta, joka kattaa koko analysoitavaksi halutun RNA:n alueen, joka on leikattu haluttuun kokoon ja johon on kiinnitetty sekvensointimenetelmän vaativa adapteri.

Sekvensointikirjaston valmistuksen ydintehtävä on näyte-RNA:n muokkaaminen niin, että se on yhteensopivaa käytetyn sekvensointimenetelmän kanssa. cDNA-kirjaston valmistuksen pääkohdat ovat seuraavat: RNA-näytteen fragmentaatio haluttuun pituuteen, RNA:n konvertoiminen kaksijuosteiseksi cDNA:ksi, oligonukleotidadapterien liittäminen templaattifragmenteihin ja kirjaston kvantitointi sekvensointia varten (Head ym. 2015).

Optimaalinen kirjaston koko riippuu käytetystä sekvensointimenetelmästä. Fragmentaatioon on käytettävissä useampia menetelmiä. Menetelmät voivat olla fysikaalisia, entsymaattisia tai kemiallisia. Useimmiten fragmentaatio tehdään joko fysikaalisesti ultraäänileikkaamalla tai entsymaattisesti käyttäen ei-spesifejä endonukleasiseoksia tai tagmentaatioreaktioita, jolloin transposaasientsyymi samanaikaisesti fragmentoi DNA:n ja insertoi adapterit kaksijuosteiseen DNA:han. Useimmiten RNA fragmentoidaan ennen konvertoimista cDNA:ksi. RNA voidaan fragmentoida digestoimalla kontrolloidussa lämpötilassa kahdenarvoisia metallikationeja (magnesium tai sinkki) sisältävässä reaktiossa. Fragmentaatio voidaan kuitenkin tehdä myös cDNA:ksi konvertoinnin jälkeen, jolloin voidaan käyttää ultraäänileikkausta (Head ym. 2015). Tässä työssä käytetyssä Lexogen Sense total RNA-seq library prep kitissä fragmentaatio tehdään käänteistranskription yhteydessä. Käänteistranskriptioon käytetään universaaleja starter- ja stopper-alukkeita, jotka kiinnittyvät satunnaisesti RNA-juosteeseen. Käänteistranskriptio alkaa starter-alukkeesta ja loppuu seuraavaan stopper-alukkeeseen, johon syntynyt juoste ligoituu. Näin syntyy satunnaisen kokoisia juosteita, joihin syntetisoidaan toinen juoste, ja tuloksena on kaksijuosteista cDNA:ta. (Lexogen.) Kirjaston kokoa voidaan säätää vielä puhdistuksen yhteydessä, jolloin adapteridimeerit poistetaan kirjastosta (Head ym. 2015).

Kaksijuosteiseen cDNA-kirjastoon lisätään adapterit, jotka määräytyvät käytetyn sekvensointialustan perusteella. Adapterit voidaan kiinnittää cDNA-juosteisiin ligoimalla tai kirjaston amplifioinnin yhteydessä. (Head ym. 2015.) Amplifioinnin yhteydessä voidaan kirjastoon vielä lisätä kullekin kirjastolle indeksi, joka on yksilöllinen usein kahdeksan nukleotidin pituinen sekvenssi, jonka avulla kukin kirjasto voidaan sekvensoida samassa sekvensointiajossa ja identifioida myöhemmin analysoinnin yhteydessä (Lexogen).

#### 2.3.4 Sekvensointikirjaston laadun varmistaminen

Sekvensointikirjaston laadun takaamiseksi on ensiarvoisen tärkeää puhdistaa kirjasto nukleotideista ja adapteridimeereistä. Adapteridimeerit syntyvät, kun adapterit ligoituvat keskenään. Nämä dimeerit muodostavat klustereita hyvin tehokkaasti pienen kokonsa vuoksi, vievät tilaa sekvensoitavilta fragmenteilta sekvensoinnin aikana tuottamatta hyödyllistä dataa. (Head ym. 2015.)

Kirjastojen puhdistukseen on useita magneettihelmien käyttöön perustuvia menetelmiä, jolloin sitoutumisolosuhteita muuttamalla voidaan vaikuttaa sitoutuvan DNA:n kokoon, ja lyhyiden dimeerien sitoutuminen magneettihelmille voidaan estää. Vaihtoehtoisesti kirjasto voidaan puhdistaa käyttötarkoitukseen tarkoitetuilla kolonneilla, agarosigeelelektroforeesilla tai pulssikenttäelektroforeesiin perustuvalla laitteistolla. Kirjaston kokojakauma voidaan tarkistaa agarosigeelelektroforeesilla tai kapillaarielektroforeesiin perustuvalla laitteistolla.

### 3 Materiaalit ja menetelmät

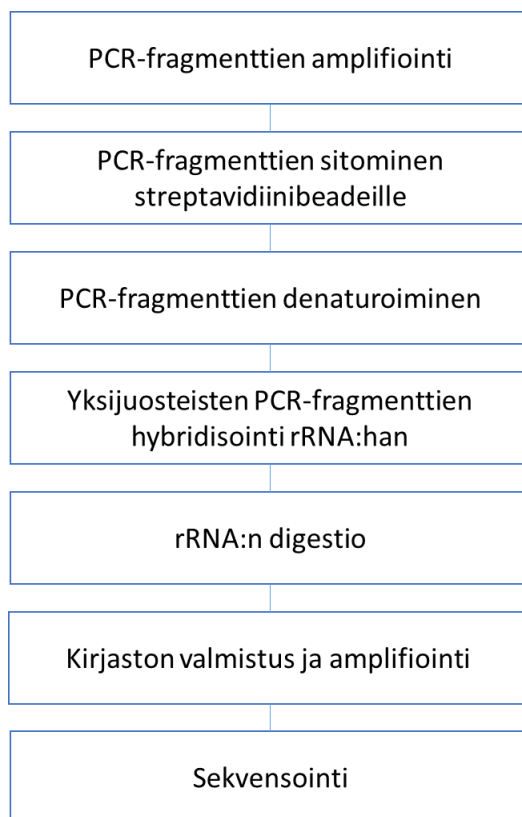
#### 3.1 Käytetyt materiaalit

Menetelmän kehityksessä käytetyt DNA:t ja RNA:t olivat seuraavista maitohappobakteerikannoista eristettyjä:

- *Leuconostoc gelidum* subsp. *gasicomitatum* KG16-1
- *Lactobacillus oligofermentans* LMG 22743
- *Lactococcus piscium* MKFS47.

### 3.2 Menetelmät

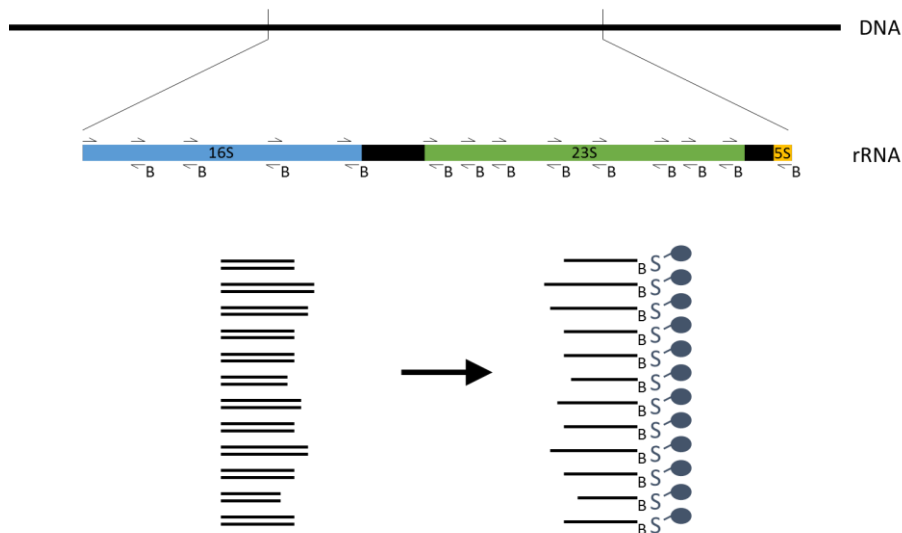
Työ koostui kuvan 1 mukaisista vaiheista. Ensimmäisenä amplifioitiin rRNA:n poistoon tarvittavat koettimet. Näitä koettimia kutsutaan jatkossa PCR-fragmenteiksi. PCR-fragmentit sidottiin streptavidinihelmille ja denaturoitiin yksijuosteisiksi. Yksijuosteiset fragmentit hybridisoitiin RNA-näytteeseen ja rRNA hajotettiin entsymaattisesti. Tämän jälkeen näytteestä valmistettiin cDNA-kirjasto ja kirjastoa amplifioitiin, jotta kirjastoa oli tarpeeksi sekvensointia varten. Lopuksi sekvensoinnin tulokset analysoitiin.



Kuva 1. Työnkulku virtauskaaviona.

#### 3.2.1 PCR-fragmenttien amplifiointi

PCR-fragmenttien valmistuksen tarkoituksena oli amplifioida noin 400–500 emäsparin pituisia fragmentteja, joilla katettiin koko rRNA-operonin alue. Kullekin bakteerille valittiin alukepareja, jotka valittiin alukepoolista siten, että syntyvät fragmentit ovat mahdollisimman vähän päällekkäisiä, mutta fragmenttien välille ei jäisi vapaita nukleotideja (Kuva 2).



Kuva 2. Esimerkkikuva, jossa tapahtuu PCR-fragmenttien amplifiointi, sitominen streptavidiniibeadeille ja muokkaus ssDNA:ksi. Kuvassa nuolimaiset viivat esittävät forward- ja reverse-alukkeita. Nämä muodostavat alukeparit siten että, ensimmäiset forward- ja reverse-alukkeet muodostavat parin ja PCR-fragmentti kattaa alukkeisen välisen alueen rRNA-operonia.

Reverse-alukkeina käytettiin biotinyloituja alukkeita, jotta myöhemmässä vaiheessa kaksijuosteiset PCR-tuotteet voitiin hajottaa yksijuosteisiksi ja vain tarvittava, rRNA:lle komplementaarinen juoste saatiin talteen. Työn aikana käytettiin useampia alukepareja PCR-fragmenttien amplifiointiin. Näistä valittiin 12 onnistunutta ja vaatimukset täyttävää fragmenttia bakteerikantaa kohti. Kullekin bakteerille valitut forward-alukkeet (Bioneer) ja reverse-alukkeet (Integrated DNA Technologies) ja niiden sekvenssit, sekä amplifioitujen fragmenttien koot on esitetty taulukossa 2.

Taulukko 2. PCR-fragmenttien valmistuksessa käytetyt alukkeet ja niiden sekvenssit, ja valmistettujen fragmenttien koot.

Näyte	F-alue	Sekvenssi	R-alue	Sekvenssi	Bp
<i>L. gasicomitatum</i>					
LG1	LG_01	AGAGTTTGATCCTGGCTCAG	LG_07 B	/5Biosg/GTAGGAGTTTGGGCCGT	400
LG28	341F1	CCTACGGGAGGCAGCAG	LP_08 B	/5Biosg/CTAATCCTGTTTGCTACCCA	450
LG3	LP_08	TGGGTAGCAACAGGATTAG	LG_19 B	/5Biosg/CTCACGACACGAGCTGAC	300
LG21	LG_19	GTCAGCTCGTGTCTGAG	LG_26 B	/5Biosg/AAAGGAGGTGATCCAGCCGCA	470
LG5	LG_26	TGCGGCTGGATCACCTCCTTT	LG_28 B	/5Biosg/CTCTGGGTCTTAGCGTACTT	600
LG33	LG_28	AAGTACGCTAAGACCCAGAG	LG_35 B	/5Biosg/CTCTGACTTCTTGTAGGCG	500
LG7	LG_34	CGTGAGGGAAAGGTGAAAAG	LO_24 B	/5Biosg/GCATTCTCACTTCTAAGCGC	800
LG23	LO_24	GCGCTTAGAAGTGAGAATGC	LG_50 B	/5Biosg/CTACCTGTGTCGGTTTGC	400
LG34	LG_50	GCAAACCGACACAGGTAG	LG_56 B	/5Biosg/CCCCTTTTTGACTCCGACT	550
LG25	LM_36	GATAGGTAGGAGCCGAAGAG	LG_60 B	/5Biosg/GTAGCTTTTATCCGTTGAGCG	340
LG36	LG_60	CGCTCAACGGATAAAAGCTAC	LG_62 B	/5Biosg/CTGTCTCAGACGTTCTGA	200
LG19	LG_62	TCAGAACGTCGTGAGACAG	LG_67 B	/5Biosg/GTGATAGAGCTTAACCTCTGTG	500
<i>L. piscium</i>					
LP1	LG_01	AGAGTTTGATCCTGGCTCAG	LG_07 B	/5Biosg/GTAGGAGTTTGGGCCGT	400
LP20	341F1	CCTACGGGAGGCAGCAG	LP_08 B	/5Biosg/CTAATCCTGTTTGCTACCCA	450
LP3	LP_08	TGGGTAGCAACAGGATTAG	LG_19 B	/5Biosg/CTCACGACACGAGCTGAC	300
LP14	LG_19	GTCAGCTCGTGTCTGAG	LG_26 B	/5Biosg/AAAGGAGGTGATCCAGCCGCA	470
LP5	LG_26	TGCGGCTGGATCACCTCCTTT	LP_20 B	/5Biosg/GCTTTCTCTTCTGTTGCTA	500
LP6	LG_27	GTGGTGGATGCCTTGGCACTAG	LG_34 B	/5Biosg/CTTTTCACCTTTCCCTCAGC	500
LP7	LG_34	CGTGAGGGAAAGGTGAAAAG	LO_24 B	/5Biosg/GCATTCTCACTTCTAAGCGC	800
LP15	LO_24	GCGCTTAGAAGTGAGAATGC	LG_50 B	/5Biosg/CTACCTGTGTCGGTTTGC	400
LP16	LG_50	GCAAACCGACACAGGTAG	LP_34 B	/5Biosg/GTCTTTCCGTCCTGTCGC	450
LP21	LG_55	TGGGCACTGTCTCAACGAG	LG_60 B	/5Biosg/GTAGCTTTTATCCGTTGAGCG	450
LP18	LG_60	CGCTCAACGGATAAAAGCTAC	LG_62 B	/5Biosg/CTGTCTCAGACGTTCTGA	200
LP19	LG_62	TCAGAACGTCGTGAGACAG	LP_40 B	/5Biosg/CAACTACTTCCGGCGTAAT	500
<i>L. oligofermentans</i>					
LO1	LG_01	AGAGTTTGATCCTGGCTCAG	LG_07 B	/5Biosg/GTAGGAGTTTGGGCCGT	400
LO20	341F1	CCTACGGGAGGCAGCAG	LP_08 B	/5Biosg/CTAATCCTGTTTGCTACCCA	450
LO3	LP_08	TGGGTAGCAACAGGATTAG	LG_19 B	/5Biosg/CTCACGACACGAGCTGAC	300
LO14	LG_19	GTCAGCTCGTGTCTGAG	LG_26 B	/5Biosg/AAAGGAGGTGATCCAGCCGCA	470
LO24	LG_26	TGCGGCTGGATCACCTCCTTT	LO_12 B	/5Biosg/CTCCTAGTACCAAGGCATTC	550
LO6	LG_27	GTGGTGGATGCCTTGGCACTAG	LG_34 B	/5Biosg/CTTTTCACCTTTCCCTCAGC	500
LO7	LG_34	CGTGAGGGAAAGGTGAAAAG	LO_24 B	/5Biosg/GCATTCTCACTTCTAAGCGC	800
LO16	LO_24	GCGCTTAGAAGTGAGAATGC	LG_50 B	/5Biosg/CTACCTGTGTCGGTTTGC	400
LO17	LG_50	GCAAACCGACACAGGTAG	LP_34 B	/5Biosg/GTCTTTCCGTCCTGTCGC	450
LO22	LG_55	TGGGCACTGTCTCAACGAG	LG_60 B	/5Biosg/GTAGCTTTTATCCGTTGAGCG	450
LO30	LG_60	CGCTCAACGGATAAAAGCTAC	LG_62 B	/5Biosg/CTGTCTCAGACGTTCTGA	200
LO31	LG_62	TCAGAACGTCGTGAGACAG	LG_67 B	/5Biosg/GTGATAGAGCTTAACCTCTGTG	500

Biotinyloidut PCR-fragmentit amplifioitiin taulukon 3 mukaisella PCR-reaktiolla. Reaktiot tehtiin Phusion Hot Start II High-Fidelity DNA Polymerase -kitillä (Thermo Scientific, tuotenro F-549L, lot 00478686). Mastermix valmistettiin ja jaettiin sekä alukkeet lisättiin reaktioihin puhdashuoneessa. Templaattina käytettiin bakteereista eristettyä DNA:ta ja fragmentit valmistettiin erikseen jokaisesta bakteerikannasta. *L. gasicomitatum* -templaatti otettiin 10-kertaisesta laimennoksesta. Templaattit lisättiin mastermixin ollessa jäällä.

Taulukko 3. PCR-fragmenttien amplifiointiin käytetty PCR-reaktio.

PCR reaktio	1	x	Pitoisuus
5x HF puskuri	10	µl	1x
10 mM dNTP:t	1	µl	200 µM
2 U/µL Phusion HS II polymeraasi	0,5	µl	1 U
DMSO 100 %	1,25	µl	2,5 %
MQ	34,25	µl	
10 µM forward-aluke	1	µl	0,2 µM
10 µM biotinyloitu reverse-aluke	1	µl	0,2 µM
Templaatti DNA	1	µl	
	50	µl	

Amplifiointiin käytettiin Veriti 96 well Thermal Cycler -laitetta (Applied Biosystems). Käytetty PCR-ohjelma on kuvattu taulukossa 4.

Taulukko 4. PCR-fragmenttien amplifiointiin käytetty PCR-ohjelma.

1.	98 °C	30 s
2.	98 °C	10 s
3.	55 °C	30 s
4.	72 °C	10 s
5.	kohtaan 2. 24 krt	
6.	72 °C	5 min
7.	4 °C	∞

Saatujen PCR-fragmenttien koot tarkistettiin agarosigeelielektroforeesilla. Ajoa varten valmistettiin geeli, joka sisälsi 1,5, % agarosia (Bioline, tuotenro BIO-41025, lot ES520-BO21900) ja 2 tippaa etidiumbromidia (Sigma, tuotenro E-8751, lot 115H 3661). Näytettä pipetoitiin geelille 5 µl ja geeliväriä 5 µl. Ladderina käytettiin 4 µl 1 kb ladderia (Gene Ruler 1 kb, Thermo Scientific, tuotenro SM0331, lot 00264162). Näytteitä ajettiin 100 V:n jännitteellä 50 minuutin ajan.



*L. gasicomitatum* kannasta amplifioitua PCR-fagmentit sekvensoitiin Abi prism 3130x1 sanger sekvensaattorilla (Applied Biosystems), jotta voitiin varmistaa, että amplifioitua fragmentit todella kattoivat koko rRNA-operonin. Sekvensoinnin tulokset analysoitiin Gap4-ohjelmassa (Staden Package).

### 3.2.2 PCR-fragmenttien puhdistus

PCR-fragmentit puhdistettiin ExoSAP-reaktiolla. Eksonukleaasi I (Thermo Scientific, tuotenro EN0582, lot 00414354) pilkkoo PCR-reaktiosta yli jääneet alukkeet ja FastAP (lämpöherkkä alkaalinen fosfataasi, Thermo, tuotenro EF0651, lot 00187985) hydrolysoi dNTP:t. Jokaiselle PCR-tuotteeseen lisättiin 10 µl ExoSAP-reaktioseosta ja reaktioita inkuboitettiin Veriti 96 well Thermal Cycler -laitteessa (taulukko 5).

Taulukko 5. ExoSAP-reaktio ja inkubointi-ohjelma

ExoSAP puhdistus 46 µl näytteelle	1	x	Inkubointi-ohjelma
Eksonukleaasi I	0,025	µl	1. 37 °C 30 min
FastAP	0,25	µl	2. 95 °C 5 min
MQ	9,725	µl	3. 4 °C ∞ (30-60 min)
	10	µl	

ExoSAP-puhdistusreaktion jälkeen kaikille fragmenteille tehtiin 9,5-prosenttinen PEG-puhdistus (engl. polyethylene glycol) tarkoituksena poistaa ylimääräinen biotiini, joka oli jäänyt seokseen ylimääräisistä alukkeista. MyOne-magneettihelmiä (Invitrogen Dynabeads MyOne Carboxylic Acid, Thermo Fisher Scientific, tuotenro 65012) otettiin 20 µl puhtaaseen 1,5 ml:n putkeen. Putki laitettiin magneetille ja neste poistettiin. Magneettihelmille lisättiin 9 µl 5 M NaCl-liuosta, 56 µl näytettä ja 40 µl PEG stock 1 -liuosta (25 % PEG, 1,5 M NaCl) jokaisen lisäyksen kohdalla hyvin sekoittaen. Reaktioseosta inkuboitettiin 10 min, jonka jälkeen supernatantti poistettiin näyteputken ollessa magneetilla. Magneettihelmet pestiin kahdesti 400 µl 80-prosenttisellä etanolilla, jonka jälkeen helmiä kuivattiin 5 min RT. PCR-fragmentit eluointiin 20 µl 1 x TE-puskuriin ja magneettihelmet eroteltiin magneetilla.

PEG-puhdistuksen jälkeen näytteiden DNA-pitoisuus määritettiin käyttäen Qubit dsDNA HS Assay -kittiä (Invitrogen, Thermo Fisher Scientific Q32854, lot 1859805,) ja Qubit 3 -fluorometriä (Invitrogen, Thermo Fisher Scientific).

### 3.2.3 PCR-fragmenttien sitominen streptavidini-magneettihelmille

Puhdistetut PCR-fragmentit sidottiin streptavidini-magneettihelmille (Dynabeads M-270, 65306, lot 167594210, Invitrogen, Life Technologies AS). Sitominen suoritettiin magneettihelmien ohjeen mukaisesti (Invitrogen). Oikean magneettihelmi- ja näyttemäärän varmistamiseksi protokolla suoritettiin ensimmäisenä fragmentille LP15, jonka jälkeen protokolla toistettiin kaikille näytteille.

Näytettä kohden pestiin 10 µl magneettihelmiä pesupuskurilla 1 x 1 ml 3 x alkutilavuus. Pestyt helmet jaettiin erillisiin putkiin, niin että niiden määrä vastasi 10 µl:n alkutilavuutta. Pesupuskuri poistettiin magneetin avulla ja helmet suspensoitiin 10 µl:aan sitomispuskuria. Tämän jälkeen helmille lisättiin näytettä tilavuus, joka sisälsi laskennallisesti 170 ng DNA:ta. Helmiä inkuboitiin 20 minuuttia hitaassa rotaatiossa RT. Supernatantti poistettiin ja helmet pestiin kahdesti pesuliuksella ja suspensoitiin 10 µl:aan sitomispuskuria. Supernatantin DNA-konsentraatio määritettiin käyttäen Qubit dsDNA HS Assay -kittiä, jotta voitiin varmistaa, että DNA oli sitoutunut helmille. Jokaisesta PCR-fragmentista otettiin 5 µl ja fragmentit yhdistettiin pooliksi.

### 3.2.4 dsDNA:n muokkaus ssDNA:ksi

PCR-fragmenttipooliin lisättiin 100 µl 0,15 M natriumhydroksidiliuosta ja seosta inkuboitiin 5 min RT. Inkuboinnin jälkeen NaOH-liuos erotettiin magneetilla magneettihelmistä ja poistettiin. Magneettihelmet pestiin kerran 100 µl 0,15 M NaOH-liuksella ja kolmesti 200 µl vedellä. Magneettihelmet suspensoitiin 100 µl 1 X TE-puskuriin ja suspension pH mitattiin pH-paperilla, jotta voitiin varmistua, että NaOH-jäämiä ei ole.

### 3.2.5 rRNA:n poistaminen RNA-näytteestä

Testi-RNA (R54 S5GPO\_2) oli eristetty kasvukokeesta, jossa oli kaikkia kolmea bakteeria kasvatettu viiden tunnin ajan. Bakteerikantojen määrien suhteita ei tiedetä, eikä näin näytteessä olevan RNA:n eri bakteerien osalta tiedetä. Näytteessä oli RNA:ta yhteensä 1052 ng/µl. Näytteestä otettiin neljä osanäytettä, joista kaksi sisälsi 1,0 µg RNA:ta ja kaksi 0,5 µg RNA:ta. Näytteistä poistettiin genominen DNA käyttäen Heat&Run-kittiä (ArcticZymes, 80200-50, lot 1731). Protokolla tehtiin kitin ohjetta mukailten. RNA-näytteen ja RNAasi vapaan veden yhteistilavuus oli 6,5 µl. Lisäksi reaktioseoksessa oli 0,65 µl 10-kertaista reaktiopuskuria ja 1 µl HL-dsDNAasia. Reak-

tioseoksia inkuboitin PCR-laitteessa (Biorad C1000) 10 min 37 °C:ssa, jonka jälkeen reaktioseoksen lämpötila nostettiin 58 °C:seen 5 minuutiksi entsyymien inaktivoimiseksi.

Genomisen DNA:n poiston jälkeen tehtiin hybridisaatioreaktio sekoittamalla 10 µl streptavidinihelmiin sidottuja PCR-fragmentteja ja 8 µl:aan RNA:ta. Reaktioseoksia inkuboitin PCR-laitteessa. Reaktioseoksia lämmitettiin 2 min 95 °C:ssa, minkä jälkeen ne hitaasti jäähdytettiin 0,1 °C/s 22 °C:seen. Tämän jälkeen reaktioita pidettiin 22 °C:ssa 5 min.

Reaktioseoksiin lisättiin 1 µl RNAasiH:ta, joka hydrolysoi RNA:n fosfodiesterisidokset, kun RNA on hybridisoituneena DNA:han. Reaktioseosta inkuboitin PCR-laitteessa 30 min 37 °C:ssa ja entsyymi inaktivoitiin 65 °C:ssa 20 min ajan.

Ribosomaalisen RNA:n digestion jälkeen PCR-fragmentit digestoitiin RapidOut DNA Removal kitillä (Thermo Scientific, lot 00541125). Digestio tehtiin kitin ohjeen mukaisesti sillä poikkeuksella, että reaktioseokseen lisättiin 0,5 µl RNAasi-inhibiittoria (40 U/µl, Ribolock RI, Thermo scientific, lot 00244718). Valmiit RNA-näytteet siirrettiin säilytykseen -80 °C:n pakastimeen.

### 3.2.6 RNA-sekvensointikirjaston valmistus

RNA-näytteistä valmistettiin sekvensointikirjastot Sense total RNA-Seq Library Prep kit for Illumina (Lexogen, lot 3959) ohjetta mukaillen (Lexogen). Käänteistranskriptioon käytettiin 8 µl RNA:ta, josta rRNA oli poistettu edellisen kappaleen menetelmän mukaisesti. RNA denaturoitiin ja reaktioseosta inkuboitin PCR-laitteessa (MJ Research PTC-200, BioRad) taulukon 6 kohtien 1 ja 2 mukaisesti. Tämän jälkeen reaktioihin lisättiin kitin entsyymiseosta ja reaktioseoksia inkuboitin taulukon 6 kohtien 3-6 mukaisesti PCR-laitteessa. Inkubointi 37 °C:ssa kesti kitin ohjeesta poiketen kaksi tuntia.

Taulukko 6. PCR-ohjelma, jota käytettiin käänteistranskriptioon ja ligaatioon.

1.	94 °C	3 min
2.	25 °C	15 min
3.	25 °C	2 min
4.	37 °C	2 h
5.	70 °C	5 min
6.	25 °C	∞

Seuraavassa vaiheessa syntetisoitiin toinen juoste cDNA:han ja hybridisoitunut RNA hydrolysoitiin. Reaktioita inkuboitiin PCR-laitteessa taulukon 7 mukaisella ohjelmalla.

Taulukko 7. Toisen juosteen syntetisointiin käytetty PCR-ohjelma.

1.	98 °C	90 s
2.	65 °C	60 s
3.	72 °C	5 min
4.	25 °C	∞

Valmiit kirjastot puhdistettiin käyttäen Lexogenin kittiä (Purification Module with Magnetic beads, Lexogen, tuotenro 022.96, lot 3887).

Puhdistettujen kirjastojen amplifioimiseksi tehtiin Lexogenin kitin ohjeen mukainen indeksi-PCR -reaktio käyttäen 5 µl valittuja i7 -indeksejä (i7 Index Plate for Quant Seq / Sense for Illumina (7001-7096), Illumina, 044,96, lot 3607). PCR-reaktio tehtiin PCR-laitteessa (Veriti, Applied Biosciences). Kitin ohjeesta poiketen alkudenaturaatio kesti 1 min ja denaturaatiovaihe kesti 30 s. Lisäksi syklejä toistettiin yhteensä 30 kertaa (Taulukko 8).

Taulukko 8. Kirjaston amplifiointiin käytetty PCR-ohjelma.

1.	98 °C	1 min
2.	98 °C	30 s
3.	65 °C	20 s
4.	72 °C	30 s
5.	kohtaan 2. 29 krt	
6.	72 °C	1 min
7.	10 °C	∞

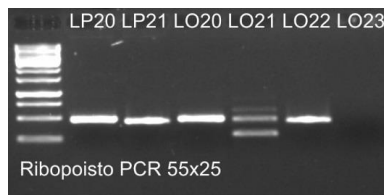
Amplifioidut kirjastot puhdistettiin 0,9x AMPurella (Agencourt), ja niiden koot määritettiin Fragment Analyzerilla. Tämän jälkeen kirjastoista tehtiin pooli niin, että kirjastojen konsentraatiot poolissa olivat yhtäsuuria. Poolille tehtiin 8-prosenttinen PEG-puhdistus, jonka jälkeen poolin puhtaus arvioitiin Fragment Analyzerilla. Koska poolissa oli jäljellä sekvensointia haittaavia n. 130 bp:n suuruisia fragmentteja, tehtiin vielä Blue Pippin -kokoselektio. Kokoselektiota varten aiemmissa puhdistuksissa syntyneet supernatantit yhdistettiin pooliin ja pooli konsentroitiin Amicon Ultra -konsentraatiopylväällä. Blue Pippinin jälkeen ajettiin Fragment Analyzer, jonka jälkeen viimeiset epäpuhtaudet pois-

tettiin 0,9x AMPure -puhdistuksella. Valmis pool tarkistettiin vielä kertaalleen Fragment Analyzerilla.

## 4 Tulokset

### 4.1 PCR-fragmenttien valmistus

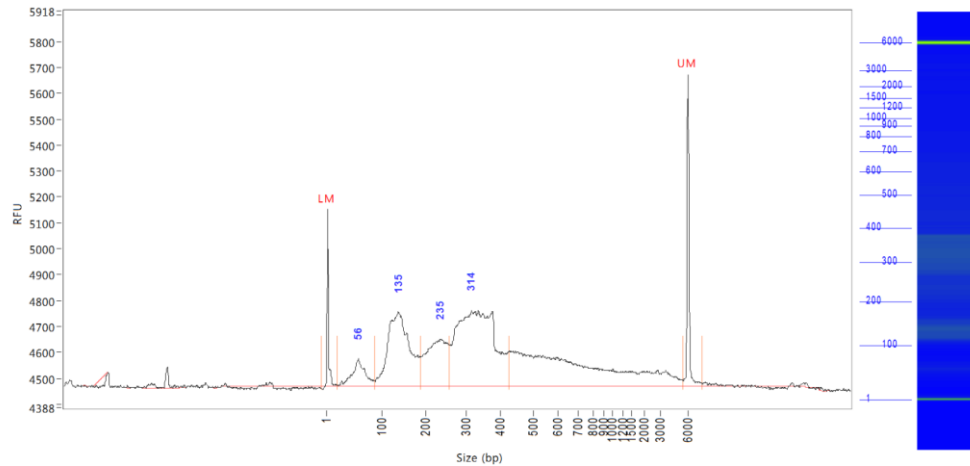
Kullekin bakteerille (*L. gasicomitatum*, *L. oligofermentans* ja *L. piscium*) saatiin valmistettua 12 PCR-fragmenttia koettimiksi. PCR-fragmentit olivat kooltaan 200–800 emäsparia. Käytetyt alukkeet ja niiden sekvenssit on esitetty taulukossa 2. Kaikki syntyneet PCR-fragmentit tarkistettiin agarosigeelielektroforeesilla. Kuvassa 3 on esimerkikuva kahdesta *L. piscium*- ja kolmesta *L. oligofermentans* -bakteerin PCR-fragmentista. Agarosigeelielektroforeesikuvat kaikista tehdyistä fagmenteista on liitteessä 1. Kuvissa on mukana useampia PCR-tuotteita kuin lopulliseen pooliin sisällytettiin.



Kuva 3. Agarosielektroforeesikuva kahdesta *L. piscium*- ja kolmesta *L. oligofermentans* -bakteerin PCR-fragmentista.

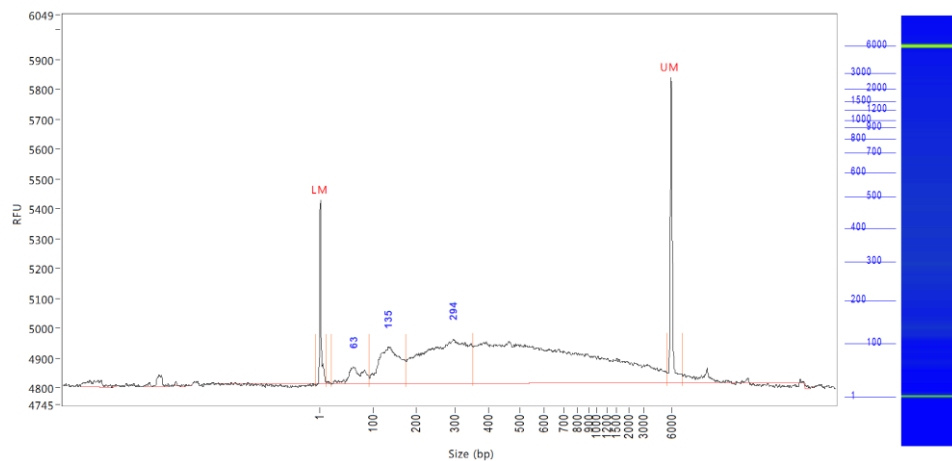
### 4.2 Sekvensointikirjaston valmistus

Sekvensointikirjastojen kokojakaumaa ei analysoitu ennen puhdistusta. Ensimmäinen puhdistus tehtiin 0,9-kertaisella AMPurella, ja sen jälkeinen kokojakauma on esitetty kuvassa 4. Kuten kuvassa nähdään, 130–140 bp kohdalla on hyvin suuri piikki. Tämä piikki johtuu dimeereistä ja haittaa suuresti sekvensointia, koska lyhyet fragmentit tarttuvat hanakammin sekvensointialukkeisiin ja näin suuria fragmentteja sekvensoituu määrällisesti vähemmän.



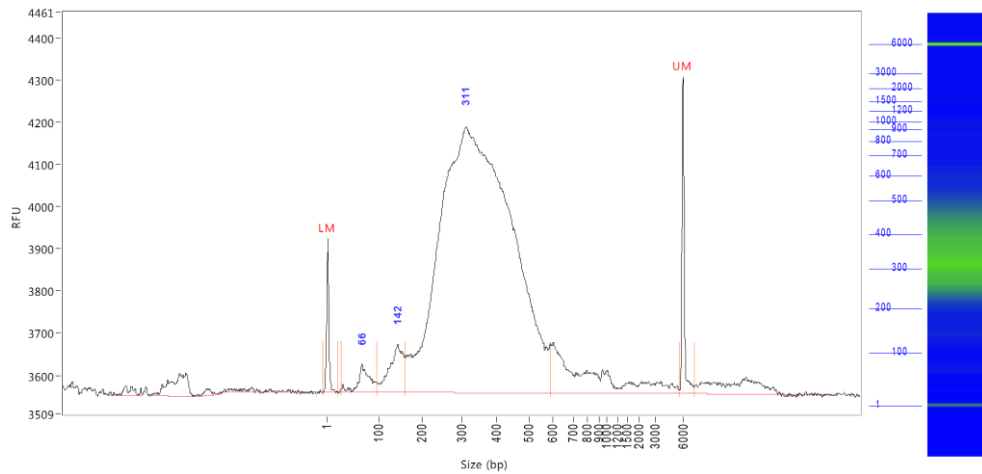
Kuva 4. Esimerkkikuva sekvensointikirjastojen kokojakaumasta AMPure-puhdistuksen jälkeen.

AMPure-puhdistuksen jälkeen näytteet poolattiin, ja PEG-puhdistettiin. PEG-puhdistus ei poistanut dimeerejä riittävässä määrin, kuten kuvasta 5 voidaan nähdä.

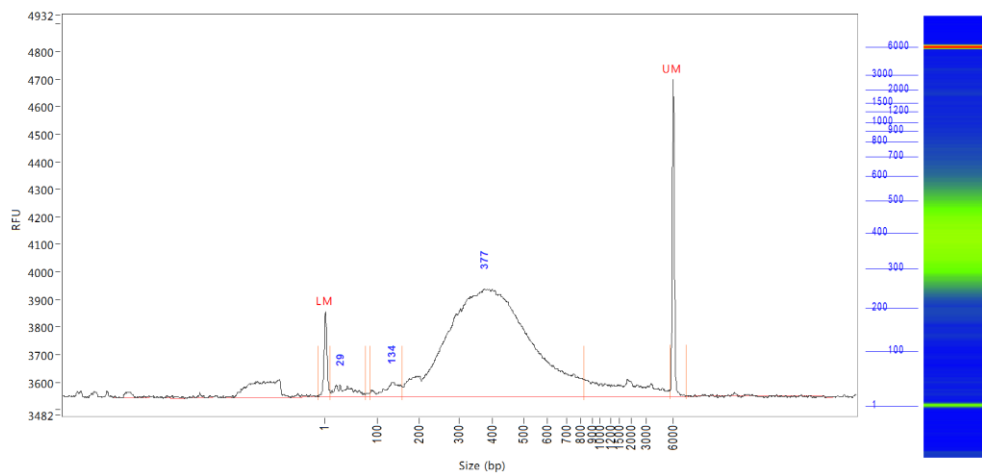


Kuva 5. Sekvensointikirjastopoolin kokojakauma PEG-puhdistuksen jälkeen. Dimeerit nähtävissä piikkinä 130 bp:n kohdalla.

PEG-puhdistuksen jälkeen supernatantit kerättiin sekvensointipoolin kanssa yhteen ja tälle poolille tehtiin Blue Pippin -kokoselektio. Kokoselektiolla saatiin poistettua suuri osa 130-140 bp:n dimeereistä (Kuva 6). Jotta sekvensointi tapahtuisi mahdollisimman optimaalisesti, pyrittiin vielä pienimmistä fragmenteista eroon tekemällä yksi AMPure-puhdistus. Tämän puhdistuksen jälkeinen kokojakauma on muuttunut (Kuva 7). AMPure-puhdistuksen jälkeen sekvensointipooli oli valmis sekvensoitavaksi. cDNA-kirjasto sekvensoitiin Illuminan MiSeq-sekvensointilaitteella.



Kuva 6. Sekvensointikirjastopoolin kokojakauma Blue pippin -kokoselektion jälkeen.



Kuva 7. Lopullisen sekvensointikirjastopoolin kokojakauma.

#### 4.3 Sekvensoinnin tulokset

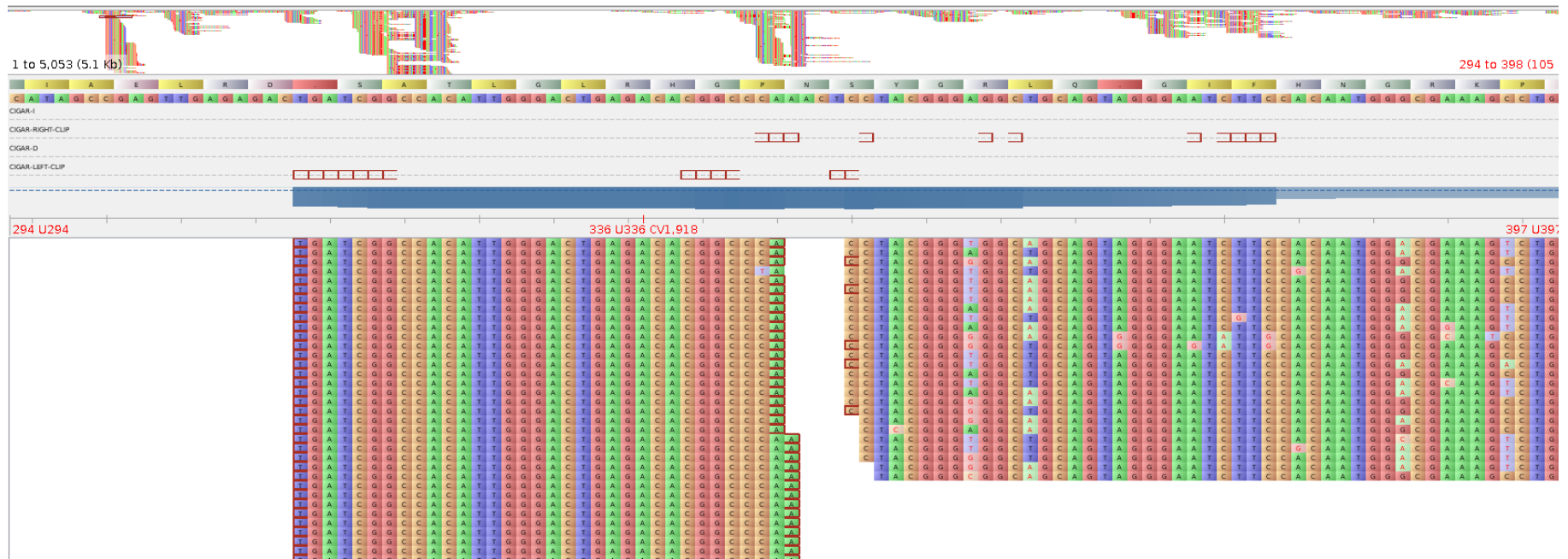
Sekvensoinnin tulokset mapattiin rRNA-operoniin, eli saatuja sekvenssejä verrattiin näytebakteerien rRNA-operonien sekvensseihin. *L. gasicomitatum* näytteiden sekvensseistä 56 %, *L. oligofermentans* näytteiden sekvensseistä 55 % ja *L. piscium* näytteiden sekvensseistä 76 % mappautui rRNA-operoniin (taulukko 9).

Taulukko 9. Näytteiden rRNA-operoniin mappautuneiden sekvenssien osuus kaikista sekvensseistä. 0-5 ja 1 viittaavat RNA:n alkumääriin 0,5 ja 1 µg. A ja B erottavat rinnakkaiset näytteet toisistaan.

Näyte	Mappautunut rRNA- operoniin (%)	
	A	B
0-5 <i>L. gasicomitatum</i>	57	59
0-5 <i>L. oligofermentans</i>	56	58
0-5 <i>L. piscium</i>	74	79
1 <i>L. gasicomitatum</i>	51	55
1 <i>L. oligofermentans</i>	50	54
1 <i>L. piscium</i>	75	76

Lexogen-kirjastosta rRNA-operoniin mappautuu yli 90 % kaikista sekvensseistä, jos rRNA:ta ei ole poistettu ennen kirjaston valmistusta. Mappaustuloksista voidaan siis päätellä, että rRNA:ta on osin saatu poistettua. Mappauskuvioita tarkasteltaessa nähtiin, että rRNA:han mappautuvat sekvenssit ovat kasautuneina yksittäisille väleille torneiksi (kuva 8). Osaan rRNA-operonista ei ole mappautunut lainkaan sekvenssejä. Tämä viittaa siihen, että näillä alueilla rRNA:n poistaminen RNAasiH:lla on onnistunut, mutta tornien kohdissa RNAasiH ei ole hydrolysoinut RNA:n fosfodiesterisidoksia.





Kuva 8. Esimerkkikuva saatujen sekvenssien mappautumisesta rRNA-operonin alueelle. Yläreunassa koko rRNA-operonin alue, josta pääkuvassa suurenneen alueelta, joka käsittää n. 100 emäsparin alueen operonista.

Seuraavaksi on käsitelty vaihtoehtoisia syitä sille, miksi tietyissä rRNA-operonin kohdissa rRNA:ta ei ole pystytty poistamaan. Kun tarkasteltiin torneiksi kasautuneita sekvenssejä, huomattiin, että tornit ovat muodostuneet PCR-fragmenttien päiden kohdille. Huomio viittaa siihen, että peräkkäisten PCR-fragmenttien päät eivät hybridisoidu RNA:han, jolloin RNAasiH ei kyseisellä kohdalla toimi. PCR-fragmenttien amplifioinnissa käytettyjen alukkeiden vuoksi fragmentit ovat osin päällekkäisiä. On siis mahdollista, että päällekkäisyys estää kummankaan fragmentin hybridisoitumisen RNA:han kyseisellä kohdalla.

Ribosomaaliseen RNA:han mappautuneet sekvenssit ovat pääosin RNA:sta peräisin. Vain yksittäisiä sekvenssejä on peräisin DNA:sta. Nämä sekvenssit ovat peräisin käytetyistä PCR-fragmenteista, ja niiden olemassaolo viittaa puutteelliseen DNAasi-käsittelyyn. Näiden sekvenssien määrä on kuitenkin siinä määrin vähäinen, että se ei edellytä muutoksia menetelmässä. Valtaosa rRNA-operoniin mappautuvista sekvensseistä on siis peräisin näytteessä olevasta rRNA:sta. RNA:ssa on sekundaarisia rakenteita, jotka muodostuvat, kun RNA emäspariutuu itsensä kanssa. Näiden rakenteiden tulee purkautua, jotta PCR-fragmentit voivat hybridisoiua täydellisesti RNA:han. Puutteellinen sekundaaristen rakenteiden purkautuminen voisi selittää sen, että tietyt kohdat RNA:sta jää näytteeseen. Menetelmässä käytettiin kuitenkin 95 °C:n lämpötilaa RNA:n denaturointiin, mikä on riittävä lämpötila sekundaaristen rakenteiden purkautumiseen, ja tämä siis tuskin on syynä jäljelle jääneeseen rRNA:han.

Morlan ym. (2012) olivat käyttäneet vastaavassa menetelmässä 50–80 emäsparin pituisia oligomeereja, mutta tässä työssä valmistetut PCR-fragmentit olivat jopa kymmenen kertaa näitä oligomeereja pitempiä. On siis mahdollista, että valmistettujen PCR-fragmenttien suuri koko vaikeuttaa DNA:n hybridisointia RNA:han. Morlanin ym. (2012) käyttämät oligomeerit eivät olleet päällekkäisiä. Niin toteutettuna menetelmä toimi, joten todennäköistä on, että PCR-fragmenttien osittainen päällekkäisyys aiheutti nyt käytetyssä menetelmässä havaitut ongelmat.

## Yhteenveto

Työn tavoitteena oli kehittää menetelmä, jolla saataisiin rRNA:ta poistetuksi RNA-näytteestä. Kehitetyllä menetelmällä RNA-näytteistä saatiin rRNA:ta osin poistettua. Menetelmä tällaisena ei kuitenkaan ole vielä käyttökelpoinen, vaan vaatii lisäkehitystä. Seuraavana vaiheena tulisi kokeilla sellaisten rRNA:lle spesifisten oligomeerien käyttöä, jotka eivät ole päällekkäisiä. Näin saataisiin varmistettua, onko PCR-fragmenttien päällekkäisyys syy ilmenneeseen ongelmaan. Oligomeerejä ei kuitenkaan lopullisessa cDNA-sekvensointikirjaston valmistamiseen tarkoitetussa menetelmässä voida käyttää, sillä rRNA:n sekvenssi tulee olla tunnettu, jotta voidaan valita käytettävät oligomeerit. Suolistomikrobinäytteiden kohdalla ei tunneta kaikkia näytteissä olevia mikrobeja saati niiden rRNA-operonin sekvenssejä.

Jos menetelmä saadaan toimimaan testissä käytetyille kolmelle bakteerille, seuraavassa vaiheessa kokeiltaisiin menetelmää suolistomikrobinäytteille. Yksittäisestä näytteestä valmistettaisiin biotinyloidut PCR-fragmentit DNA-näytteestä universaaleja alukkeita käyttäen, jolloin saadaan kaikkia näytteessä olevia bakteereja vastaavia PCR-fragmentteja. Universaaleja alukkeita käytettäessä tosin tulee pitää huolta, ettei synny päällekkäisyyksiä. Tämän jälkeen näitä PCR-fragmentteja voidaan käyttää jokaisen RNA-näytteen kohdalla, ja rRNA:n poisto ja RNA-sekvensointikirjaston valmistaminen voidaan tehdä, kuten tässä työssä on kuvattu.

## Lähteet

Agencourt. Agencourt AMPure PCR purification. Ohje 000601v024.

He, Shaomei; Wurtzel, Omri; Singh, Kanwar; Froula, Jeff; Yilmaz, Suzan; Tringe, Susannah; Wang, Zhong; Chen, Feng; Lindquist, Erika; Sorek, Rotem & Hugenholtz, Philip. 2010. Validation of two ribosomal RNA removal methods for microbial meta-transcriptomics. *Nature Methods*. Vol. 7, s. 807–812.

Head, Steven; Komori, H Kiyomi; LaMere, Sarah; Whisenant, Thomas; Van Nieuwerburgh, Filip; Salomon, Daniel & Ordoukhanian, Philip. 2015. Library construction for next-generation sequencing: Overviews and challenges. *Biotechniques*. Vol. 56, s. 61–91.

Lexogen. Sense total RNA-seq library prep kit – user guide, 009UG013V0141. Itävalta: Lexogen.

Mardis, Elaine. 2017. DNA sequencing technologies: 2006–2016. *Nature Protocols*. Vol. 12, s. 213–218.

Morlan, John; Qu, Kunbin & Sinicropi, Dominick. 2012. Selective depletion of rRNA enables whole transcriptome profiling of archival fixe tissue. *PLOS ONE*. Vol. 7, s. 1–8.

Pereira, Pedro; Aho, Velma; Paulin, Lars; Pekkonen, Eero; Auvinen, Petri & Scheperjans, Filip. 2017. Oral and nasal microbiota in Parkinson's disease. *Parkinsonism & Related Disorders*. Vol. 38, s. 61–67.

Petrova, Olga; Garcia-Alcalde, Fernando; Zampaloni, Claudia & Sauer, Karin. 2017. Comparative evaluation of rRNA depletion procedures for the improved analysis of bacterial biofilm and mixed pathogen culture transcriptomes. *Scientific reports*. Vol. 7, artikkeli 41114.

Scheperjans, Filip; Aho, Velma; Pereira, Pedro; Koskinen, Kaisa; Paulin, Lars; Pekkonen, Eero; Haapaniemi, Elena; Kaakkola, Seppo; Eerola-Rautio, Johanna; Pohja, Marjatta; Kinnunen, Esko; Murros, Kari & Auvinen, Petri. 2015. Gut microbiota are related to Parkinson's disease and clinical phenotype. *Movement Disorders*. Vol. 30, s. 350–358.

Stranneheim, Henrik & Lundeborg, Joakim. 2012. Stepping stones in DNA sequencing. *Biotechnology Journal*. Vol. 7, s. 1063–1073.

Suominen, Ilari; Pärssinen, Raimo; Haajanen, Kari & Pelkonen, Jani. 2010. *Geeniteknikka*. Turku: Turun ammattikorkeakoulun oppimateriaaleja 52.

## PCR-fragmenttien agarosigeelielektroforeesikuvat

